How the Nature of Expression Constitutes a Problem for Expressivism

James Sias

A thesis submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Master of Arts in the Department of Philosophy.

Chapel Hill
2009

Approved by

Dorit Bar-On (Chair)

Geoffrey Sayre-McCord

Keith Simmons

ABSTRACT

JAMES SIAS: How the Nature of Expression Constitutes a Problem for Expressivism
(Under the direction of Dorit Bar-On)

Expressivism in ethics is the view that moral language functions to express only the conative mental states of speakers. After distinguishing between two different senses of expression, and three different ways in which the expressivist might account for the expressive function of moral language, I explain that expressivists owe us an account of how it is that speakers express their mental states by making moral claims. Then I go on to develop just such an account, and in the end, the account puts expressivists in an awkward position. For if my account of expression is on point, then either (a) expressivists are committed to saying that competent users of moral language are often confused (perhaps even systematically so) about their own mental states and their own reasons for action, or (b) expressivism is false.

To James.
Hooray to him!

TABLE OF CONTENTS

1. INTRODUCTION

Let me start by setting the scene. Much, if not every last bit, of the work done in metaethics finds its roots in G. E. Moore's exposure of an apparent tension between philosophical naturalism, on the one hand, and our ordinary, commonsense notions of moral goodness, evil, and the like, on the other. In fact, if I were in the business of making flowcharts, the very first distinction I might draw among the many positions taken in metaethics would be between (a) those who agree with Moore in thinking that there is a *real* tension between naturalism and moral properties (even if there might be better ways of capturing the tension than Moore's famous open question argument) and (b) those who disagree with Moore and insist that, whatever tension there may seem to be, it is only an *apparent* one. Philosophers who take this latter position are, for the most part, united in their interests in defending some variety of what has come to be called ethical naturalism. Philosophers who take the former position, however, are far from united, for some see the tension between naturalism and moral properties to be an indictment of naturalism, while others see the tension as an indictment of moral properties. One person's modus ponens is, after all, another's modus tollens.

Now, among those who see the tension as an indictment of moral properties, we can add to our flowchart a further distinction, according to their response to the following question: If there are not any moral properties for moral terms to refer to, then what is it that we are doing when we use moral language? J. L. Mackie, for instance, denied the reality of moral properties, but he shared with moral realists the assumption that, when we use moral language, we often do so in order to express our beliefs about the moral properties of things. This, of course, led him to the conclusion that all of our moral thought and discourse rests on a mistake. Others,

however, find Mackie's error theory to be no moral palatable than the moral realism they join him in rejecting (after all, the commitment of all of humanity to such a deep, systematic error may not seem any easier to swallow than a realm of supposedly queer moral properties), and so they've had to tell a different story altogether about the function of moral language. And that story, for the most part, has been this: when we use moral language, we do not express beliefs at all; instead, we express only non-cognitive states of mind like emotions, desires, attitudes, and the like.[1] This is the view that I will be calling *expressivism*.

Expressivism certainly has its virtues. For starters, there really cannot be any question about the compatibility of expressivism with naturalism. As Simon Blackburn puts it, expressivism is a view that "intends to ask no more from the world than what we know is there," by positing "no more than this: a natural world, and patterns of reaction to it."[2] Expressivism also has the virtue of being able to give a rather easy account of the connection between sincere moral claims and motivation. It is often thought to be a conceptual truth that a person cannot sincerely claim that φing is good without having some desire or motivation to φ, or to promote φing, or to praise others who φ, etc. But if beliefs are all that is needed in order to make sincere moral claims, then this tie between moral claims and motivation is difficult to explain, since beliefs are not generally thought of as motivational states of mind. In order to best account for the tie between sincere moral claims and our motivation to act accordingly, some argue, we ought to understand the sincerity of one's claim that φing is good to consist in one's having

---

[1] It should be noted that the plausibility of one's expressivism may depend in part upon exactly what *sort* of non-cognitive mental state one suggests is expressed when we use moral language. And of course, expressivists go in a number of different directions on this point. But for my purposes, all that matters is that we understand expressivism in general as the view that, when we use moral language, it is some sort of *conative* mental state that gets expressed.

[2] Blackburn 1984: 182.

some conative state of mind like a desire or a pro-attitude toward φing, which is precisely the sort of line that expressivists take.

Along with expressivism's virtues, though, come several vices. Perhaps the most famous of expressivism's problems is the Frege-Geach problem. In his paper "Assertion," Peter Geach presses what he calls "the Frege point," which is roughly that the semantic content of a sentence ought to remain constant across both asserted and unasserted contexts.[3] This is a problem for expressivists because their theory only tells us how a moral sentence gets its content when it is asserted. In other words, if the content of a moral sentence consists in the attitude that speakers express when they are asserting that sentence, then how does a moral sentence retain its content in unasserted contexts, such as the antecedent of a conditional? For the past forty years or so, the Frege-Geach problem has been the primary point of concern in critical discussions of expressivism, but it is certainly not the only one. Frank Jackson and Philip Pettit, for instance, argue that expressivism cannot really be distinguished from subjectivism.[4] And Cian Dorr argues that expressivism allows us infer things about the way the world is on the basis of only our attitudes, an epistemic no-no that he calls "wishful thinking."[5]

In this paper, I will raise a new problem for expressivism. I'll begin in section 2 by distinguishing between three distinct kinds of expressivist views, so that I might home in on exactly the sort of view that I will take to be my primary target. In section 3, I'll take a closer look at how it is that moral language is supposed to have the sort of expressive function that expressivists claim it has. Then, in section 4, I'll begin to explore in some detail an issue that has only very recently been getting the sort of attention that I think it deserves—namely, the issue of

---

[3] Geach 1965: 449.

[4] Jackson and Pettit 1998.

[5] Dorr 2002.

3

how it is that speakers express their mental states. I want to get my hands on what it is, exactly, that expression is supposed to be for the expressivist. This will involve discussions of whether or not speakers can express—in the relevant sense of expression—mental states that they do not have; whether or not a speaker's ability to express a mental state depends upon her behaving in a way that others would recognize as an expression of that state; and whether or not expression is an intentional phenomenon. But my discussion of the nature of expression will not be complete until section 5, where I will put forth and defend a proposal for the sort of relationship that a speaker must stand in to her speech acts and her mental states in order for her to express those mental states by performing those acts. This proposal, though, will turn out to be a problem for expressivism, for I will argue that, if the proposal is on point, then either (a) the expressivist is committed to saying that competent users of moral language are often confused (perhaps even systematically so) about their own mental states and their own reasons for action, or (b) expressivism is false. And finally, in section 6, I'll consider a few ways in which expressivists might respond.

2. THREE KINDS OF EXPRESSIVISM

In this section, I'll distinguish between what I take to be three distinct kinds of expressivism. But first, I think it will prove important for me to shear apart two views that are often combined into one: *expressivism* and *non-cognitivism*. There are two issues at stake here: on the one hand, there is the issue of what sorts of mental states get expressed when we make moral claims; and on the other hand, there is the issue of whether or not moral claims are of a sort that admit of truth or falsity. And oftentimes, when either expressivism or non-cognitivism is described, it is described as a view that involves both of these issues at once. Take, for example, Jackson and Pettit's description of expressivism:

> Expressivism is a bipartite theory. It holds, first, that ethical sentences lack truth conditions—they are not truth-apt, truth assessable, etc.—and do not serve to report anything that the speaker believes to be so. And it holds, second, that ethical sentences express certain distinctive pro and con *attitudes*.[1]

And compare that to Russ Shafer-Landau's description of non-cognitivism:

> According to non-cognitivism, … moral judgments are not beliefs and thus are not truth-evaluable. Moral judgments … serve as expressive vehicles, primarily giving vent to our emotions, prescribing courses of conduct, or expressing our non-cognitive commitments. As such, they aren't the sort of things fit to be considered either true or false.[2]

Now, to be fair, it is easy to see why expressivism and non-cognitivism are so often combined into one. The assumption that neither an attitude nor the expression of an attitude could ever be true or false is a natural one to make. But, as I'll soon explain, there are some who reject this

---

[1] Jackson and Pettit 1998: 239, italics in original.

[2] Shafer-Landau 2003: 18. Shafer-Landau actually goes on to say that "expressivism" and "non-cognitivism" are just different names for the same view (19).

assumption. So, for my purposes, I'll be taking expressivism to *only* be a view about what sorts of mental states get expressed when speakers make moral claims; and I'll be taking non-cognitivism to *only* be a view about the truth-aptness of moral claims.

Having gotten that out of the way, I can now distinguish between three different kinds of expressivism. Let's begin with what I will call the EXPRESSIVIST THESIS:

> EXPRESSIVIST THESIS: The function of moral language is to express only the conative attitudes of speakers.

In *Language, Truth, and Logic*, we find A. J. Ayer defending something like the EXPRESSIVIST THESIS. He agrees with Moore in thinking that moral concepts are unanalyzable, but whereas Moore takes this to be an indictment of naturalism, Ayer takes it to be an indictment of moral concepts. "[T]he reason why they are unanalysable," Ayer explains,

> is that they are mere pseudo-concepts. The presence of an ethical symbol in a proposition adds nothing to its factual content. Thus if I say to someone, "You acted wrongly in stealing that money," I am not stating anything more than if I had simply said, "You stole that money." In adding that this action is wrong I am not making any further statement about it. I am simply evincing my moral disapproval of it. It is as if I had said, "You stole that money," in a peculiar tone of horror, or written it with the addition of some special exclamation marks. The tone, or the exclamation marks, adds nothing to the literal meaning of the sentence. It merely serves to show that the expression of it is attended by certain feelings in the speaker.[3]

It is not that the term "wrongly" is utterly superfluous; it's just that its function is nothing other than that of allowing a speaker to express his attitude toward his hearer's action. Consider the following two claims:

1. "You acted wrongly in stealing that money."

2. "You acted quickly in stealing that money."

In making the second claim, you might think, a speaker straightforwardly expresses the belief that his hearer's action had a certain property—namely, the property of having been performed

---

[3] Ayer 1952: 107.

quickly. But the same treatment cannot be given to (1), since "wrongly" only serves to give voice to the speaker's disapproval of the action. The speaker might just as well have said "Boo to stealing the money!" or "Thief!" in a particularly disapproving tone of voice.

Most defenders of the EXPRESSIVIST THESIS are also non-cognitivists, for exactly the reason I mentioned earlier. If moral claims amount to nothing more than expressions of emotions, or desires, or the like, then it is hard to see how they could be truth-apt, since conative states such as these are not generally thought of as truth-evaluable. My *belief* that you stole the money is truth-apt, and so is the *sentence* "You stole the money." But my *disapproval* of your thievery, it is natural to think, is not truth-apt. Call any view that conjoins the EXPRESSIVIST THESIS with non-cognitivism a form of *pure expressivism*.

There may, however, be reasons for the expressivist to resist non-cognitivism. For starters, as Blackburn has acknowledged, there doesn't seem to be anything wrong with responding to moral claims by saying, "That's true," or "That's not true," whereas these would surely seem inappropriate in response to a "Boo!" or to a mere tone of voice.[4] More significant, though, is the famous Frege-Geach problem. If moral claims are to retain their contents in such unasserted—and, importantly, truth-functional—contexts as the antecedents of conditionals, then we should expect their contents to be of a truth-apt sort.[5] For these and other reasons, some expressivists try to eschew non-cognitivism without compromising what is central to expressivism—namely, that moral language functions to express conative, and not cognitive, states of mind.

---

[4] Blackburn 1993: 111. Of course, neither Blackburn nor I take this point alone to be a particularly devastating blow to non-cognitivism. It is, rather, just *among* the things that we might not expect to be the case if non-cognitivism is true.

[5] For more on the Frege-Geach problem, see Schroeder 2008c. As a detailed explanation of the Frege-Geach problem is all but irrelevant to my aims in this paper, I'll not be offering one here.

Blackburn, for instance, defends a second kind of expressivist view that I will call *deflationary expressivism*, which begins with the EXPRESSIVIST THESIS and then "tries to earn, on the slender basis, the features of moral language … which tempt people to realism."[6] There are two ways in which the deflationary expressivist might earn the right to talk of "moral truth," one more ambitious than the other. The ambitious way is to construct a substantive property of truth that can be ascribed to moral claims without committing us to an ontology replete with moral facts and properties. Blackburn explores the possibility and virtues of such a construction in his earlier work.[7] The idea is this: each of us has a set of moral attitudes, and some sets are better than others, in the sense that they are more consistent, more conducive to the fulfillment of various practical purposes, etc. (e.g., a set that contains both a disapproval of lying and an approval of getting one's little brother to lie is an inconsistent set). So we can define a "best possible set of attitudes" as the one that is maximally consistent, etc., and just stipulate that a moral claim is *true* if it expresses one of the attitudes contained in the best possible set.[8] A less ambitious way to earn the right to talk of moral truth, though, is just to deflate the truth predicate. When we say of a moral claim that it is true, we are not ascribing any substantive sort of property to the claim at all; rather, we are just endorsing the sort of attitude that one might express by making that particular claim.[9] This is the sort of line that Blackburn takes in his later

---

[6] Blackburn 1984: 171. I confess to not really understanding what Blackburn means by "on the slender basis," but I imagine that it has something to do with the basis of deflated concepts upon which he will eventually claim that expressivists can, just as readily as realists, talk of the truth or falsity of moral claims.

[7] Blackburn 1984: 197-202.

[8] Blackburn 1984: 198. Blackburn himself acknowledges, "Although this is the simplest projectivist account of truth, and is one used by many anti-realists, I do not myself think it is the best. It is only a first approximation, but serves to make the immediate points" (1984: 198, n. 10).

[9] I should note that this is just *one* way in which the expressivist might deflate the truth predicate. Another is to take the standard disquotationalist route and say that, when Peter says, "It is true that murder is wrong," he is saying nothing different than just "Murder is wrong." And in this case, the first claim would not be an endorsement of the other, since the two are really one and the same.

work. "To think that a moral proposition is true," he explains, "is to concur in an attitude to its subject."[10] When Peter says, "It is true that murder is wrong," he is doing nothing different than if he'd said "Amen" in response to someone else saying that murder is wrong. And so, contrary to the common assumption that expressivism and non-cognitivism are attached at the hip, the deflationary expressivist claims to have just as much a right as moral realists to talk of the truth or falsity of moral claims, as long as those predicates are given this deflationary treatment.[11]

And finally, there are the *hybrid expressivisms*. The hybrid expressivist, like the deflationary expressivist, attempts to combine expressivism and cognitivism, but without having to make the move to deflationism. The idea, as Mark Schroeder explains, is this:

> [I]f cognitivists avoid expressivists' problems with logic and inference because they associate moral sentences with ordinary descriptive contents, and expressivists can offer elegant explanations of the motivating power of moral judgments … , then maybe a view according to which moral sentences express *both* kinds of states of mind could claim *both* of these kinds of advantages.[12]

One such hybrid view is David Copp's so-called "realist-expressivism."[13] Copp asks us to consider a sentence like "Your cur howled all night," said by someone to his neighbor. Strictly speaking, the word "cur" refers to a type of dog, but, as a matter of convention, it is only used when the speaker has a particularly negative attitude toward the animal of which he is speaking.

---

[10] Blackburn 1993: 129.

[11] Of course, one might wonder whether or not the deflation of the truth predicate really earns the deflationary expressivist a genuine form of cognitivism. Cognitivism, after all, is not technically a thesis about the legitimacy of our use of the truth predicate; rather, it is a thesis about the truth-aptness of moral claims (where truth-aptness is supposed to name a relationship that language bears to the world, independent of our practices of using that language). So to say that I am free to say things like "It is true that p" is not necessarily to say that p is truth-apt. I owe this point to Dorit Bar-On, though I'll not be pursuing it any further than I've done here. For more on the relationship between expressivism and deflationism, see Smith 1994, Divers and Miller 1994, and Marino 2005.

[12] Schroeder forthcoming: 1.

[13] Copp 2001.

So the sentence expresses both (a) the belief that the dog being spoken of howled all night, and, because of the conventions associated with the word "cur," (b) a negative toward the dog. Similarly, Copp argues, the claim "Murder is wrong" expresses both (a) the belief that murder is wrong, and, because of various conventions associated with moral language, (b) a negative attitude toward murder.

If Copp's realist-expressivism doesn't appeal to you, though, there are other hybrid views from which to choose. Dan Boisvert, for instance, defends a hybrid theory that he calls "expressive-assertivism," which he considers to be "a refined improvement of the 'dual-use' expressivist theories traditionally associated with C. L. Stevenson's emotivism and R. M. Hare's prescriptivism."[14] On Boisvert's theory, two distinct illocutionary acts might be performed in the performance of just one locutionary act. So, for instance, according to his Dual-Use Principle, "If a speaker correctly and literally utters a basic ethical sentence, then the speaker performs one direct *expressive* illocutionary act and one direct *assertive* illocutionary act."[15] In this way, Boisvert is supposed to have captured both the expressivist intuition (that moral claims are expressions of conative states) and the cognitivist intuition (that moral claims are assertions of beliefs, which are truth-apt). Yet another hybrid expressivist view is Dorit Bar-On and Matthew Chrisman's "ethical neo-expressivism."[16] One of the key points of ethical neo-expressivism is the distinction between *the act of making a moral claim*, on the one hand, and *the claim itself*, on the other. The latter is a declarative sentence, and as such, it expresses (in what they call the *semantic* sense of expression) a proposition; the former is an action, and as such, it expresses (in what they call the *action* sense of expression) a state of mind. So when I claim that murder is wrong,

---

[14] Boisvert 2008: 170.

[15] Boisvert 2008: 171.

[16] Bar-On and Chrisman forthcoming.

both (a) a negative attitude toward murder and (b) the proposition *that murder is wrong* (which is truth-apt) get expressed, just in two different senses of expression.

So now that I've superficially mapped the terrain of expressivist views, let me try to make as clear as I can exactly what it is that I will take to be the *target* of my criticism in this paper. As a realist, I am primarily interested in expressivism insofar as it constitutes a rival to my preferred account of moral language, according to which speakers (at least sometimes) express beliefs of theirs about putative moral facts of the matter when they use moral language.[17] This being the case, I will take pure and deflationary versions of expressivism to be closer to the crosshairs of my attack than any of the hybrid views, since neither pure expressivists nor deflationary expressivists allow for a speaker to ever express a cognitive mental state like belief when he or she uses moral language. So from this point forward, I am interested in whatever it is that pure and deflationary expressivisms have in common, and I'll understand this to be the fact that both views straightforwardly affirm the EXPRESSIVIST THESIS.[18] So it is to an analysis of the EXPRESSIVIST THESIS that I shall now turn.

---

[17] It should be noted—and it will come up again later—that moral realism is probably best understood as a view about only the *metaphysics* of morality, and as such, it is a view that is strictly speaking neutral with respect to what sorts of mental states get expressed by speakers when they use moral language. One might, for instance, hold that the world is replete with moral facts and properties, and yet also hold that, when speakers use moral language, they do not thereby express beliefs of theirs about those moral facts and properties. Traditionally, though, realists supplement their metaphysics with both (a) a rejection of expressivism and (b) a rejection of non-cognitivism. And I join them in doing so. So hereafter, when I speak of moral realism, I'll have this traditional sort of realism in mind.

[18] Hybrid expressivisms sometimes involve something of a departure from the EXPRESSIVIST THESIS. Copp's realist-expressivism and Boisvert's expressive-assertivism, for instance, both allow that moral language can function to express a speaker's attitudes *and* to express a speaker's beliefs, whereas the EXPRESSIVIST THESIS allows for only the former to be the case. This may not be the case, though, with Bar-On and Chrisman's ethical neo-expressivism. After all, since they allow that, in the *action* sense of expression, moral language can function to allow speakers to express only conative attitudes, there isn't as clear a conflict between their view and the EXPRESSIVIST THESIS as there is in the case of the other two hybrid views.

3. HOW TO ACCOUNT FOR THE EXPRESSIVE FUNCTION OF MORAL LANGUAGE

According to the EXPRESSIVIST THESIS, to claim that murder is wrong is to express only a (negative) conative attitude toward murder. But it's not yet clear what this is supposed to mean. After all, not only are there are at least two different senses of expression that might be at work in the EXPRESSIVIST THESIS, but there is also more than one story that we might tell about how it is that moral language functions to express (in whatever sense) only conative attitudes. In this section, I'll try to clear these issues up.

Recall the distinction that neo-expressivists Bar-On and Chrisman make between the sort of expression that speakers do by performing certain actions—which they call expression in the *action* sense, or a-expression—and the sort of expression that linguistic units, like sentences, do—which they call expression in the *semantic* sense, or s-expression.[1] The former is a relation between speakers and their mental states; the latter is a relation between linguistic units and their contents. So the sentence "The cat is on the mat" s-expresses the proposition *that the cat is on the mat*. But this is not to say that, when I utter that sentence, I am a-expressing any belief of mine about the cat being on the mat. After all, I may not even have such a belief. Claiming that the cat is on the mat might be my secret way of alerting a friend to some imminent danger. So with the distinction between a-expression and s-expression in hand, we might wonder which of the two senses of expression is at work in the EXPRESSIVIST THESIS? That is, when the EXPRESSIVIST THESIS says that moral language functions to express only conative attitudes,

---

[1] Bar-On and Chrisman forthcoming: 4ff. Here they borrow from Sellars (1969) who actually distinguishes these two senses of expression from a third sense—the *causal* sense of expression—according to which an utterance or piece of behavior expresses a mental state by being the end of a causal sequence that began with that mental state. This is the sense in which an unintentional wince may express a person's pain. I'll join Bar-On and Chrisman in setting this causal sense of expression to the side for the sake of my own discussion.

should that be taken to mean that (a) when speakers make moral claims, they a-express only their conative attitudes toward things, or that (b) the moral claims themselves s-express contents that are somehow or other related only to conative attitudes? This question is significant enough, I think, to deserve its own name. So I'll call it "the expression question."

For now, I'll assume that the expressivist *does* have an answer to the expression question, but I should note that the answer is not made as clear as one might hope by the literature on expressivism. Sometimes, for instance, expressivism is presented in a way that suggests that it is first and foremost a view about what it is that speakers (either typically or necessarily) a-express when they use moral language. This seems like the natural way to understand Ayer when he suggests, for instance, that when I say, "You acted wrongly in stealing the money," I might just as well have said, "You stole the money" in a disapproving tone of voice. The moral term "wrongly" is no more significant to the semantic content of my claim than is the tone of voice in which I make it. Instead, the significance of the term "wrongly" lies in its allowing me to a-express my disapproval of someone's thievery. Other times, though, expressivism is presented as first and foremost a view about what sorts of semantic contents get s-expressed by moral claims. Schroeder, for instance, treats expressivism as chiefly "a semantic program,"[2] and Ralph Wedgwood describes it as a view about "the meaning of normative statements, and of the sentences that are used to make those statements."[3] And still others present expressivism in a way that makes it unclear which of the two senses of expression really lies at the heart of the EXPRESSIVIST THESIS. Jackson and Pettit, you will recall from the quotation above, describe expressivism as a view according to which "ethical sentences express certain distinctive pro and

---

[2] Schroeder 2008a.

[3] Wedgwood 2007: 35. For the record, Wedgwood defines "statement" as "the speech act that is performed by the sincere utterance of a declarative sentence" (35).

con *attitudes*."[4] On the plausible assumption that declarative sentences are only in the business of s-expressing *propositions*, though, I think it is fair to wonder what it could mean for a declarative moral sentence to express a conative attitude. A sentence might s-express an attitude, for instance, by *referring* to the attitude, as in the sentence "I disapprove of murder." But the sentence "Murder is wrong" refers to no such attitude. So maybe the sense in which moral sentences express conative attitudes is that of being used by speakers to a-express their conative attitudes. So, assuming Jackson and Pettit are not confused in any way, their formulation of the EXPRESSIVIST THESIS might nonetheless be seen as papering over the distinction between s-expression and a-expression. Either way, my point here is just that the literature is not as helpful as we might hope when it comes to discerning the expressivist's answer to the expression question.

So then how *should* the expressivist answer the expression question? Well, I think there are actually three different routes that the expressivist might take at this point. The first is to tell a kind of diachronic story about how the semantic contents that moral claims now s-express have their origin in (and, thanks to the establishment of convention, are still tied down to) the sorts of mental states that speakers a-expressed when they first began using moral language. At some point in the past, people realized a need to communicate to others their approval or disapproval of various actions, character traits, ways of living, etc., in order to relate to each other in a cooperative, mutually advantageous manner. So moral language emerged out of a need for there to be some linguistic means for people to express their attitudes of approval, disapproval, and the like. And over time, it caught on. Moral language began to serve a rather vital social role, and in order for it to continue to do so, it had to be used with a certain measure of regularity. Regularity became convention, and it is this convention that allows moral language

---

[4] Jackson and Pettit 1998: 239, italics in original.

to have the stable semantic content that it now has. What that content *is*, though, is to be provided by the sorts of attitudes that speakers first used the language to express. According to this diachronic story, the sentence "Murder is wrong," despite its declarative form, is actually synonymous with something like "Boo to murder!" because "wrong" is a term that was originally used by speakers to a-express their disapproval of things. So the answer to the expression question, were the expressivist to take this route, is that the function of moral language is to be understood primarily in terms of its ability to s-express contents that are genetically related to speakers' conative attitudes.[5]

This certainly seems to be a route that many expressivists are inclined to take. Blackburn and Gibbard, for instance, both tell diachronic stories about how moral semantics relates back to the genesis of moral language.[6] But there are problems with the diachronic story. First, to whatever extent the expressivist's semantic program *depends* upon this being the correct story of how moral language came to be, the expressivist leaves himself open to the charge that he's given us nothing more than a just-so story that is neither verifiable nor falsifiable. We might just as well imagine moral language coming about in order to fill an explanatory need: people came to recognize that certain actions, character traits, ways of living, etc., had properties for which there were not yet any names, and so terms like "good" and "evil," "right" and "wrong" were coined so that we might be able to offer a more complete explanation of the world around us.[7] This, I take it, is how many other sorts of discourse came to be. We couldn't very well give a

---

[5] In their paper, Bar-On and Chrisman tell a similar story about how expressive language can come to have declarative, propositional form. "[U]tterances of 'Yuck!'" they write, "typically serve to a-express an utterer's state of disgust. 'Yuck,' it seems, is not a descriptive term whose function is to denote an observed property of things. For all that, 'Yuck' can be replaced by utterances that have propositional form: 'This is yucky!'" (23-24). The idea is that moral sentences like "Murder is wrong" may have come upon their declarative, propositional form in a similar way.

[6] See Blackburn 1998: ch. 3 and Gibbard 1990: ch. 3.

[7] See Sturgeon 1992.

complete account of reality without, for instance, names for colors, points in space and time, etc., and so presumably these terms came into being for largely explanatory purposes. My point, though, is not simply that we *can* imagine telling such a story about the origins of moral language. Rather, my point is that, on this diachronic story, whether or not the expressivist is telling the right story about what moral claims now s-express depends entirely upon whether or not he's told us the right story about how moral language came about. And this strikes me as a rather shaky ground upon which to rest such a substantive semantic program.

Here is another problem for the diachronic story. Technically, if expressivism is to be understood as the view that moral claims now s-express contents that draw all of their semantic significance from the conative states that speakers used them to a-express prior to the establishment of any conventional regularity, then the mental states of speakers *now* are all but irrelevant. Now that convention has already tied the meanings of moral claims down to whatever conative states those claims were originally used to a-express, it really doesn't matter what sorts of mental states we now use these claims to a-express. In fact, it is perfectly consistent with this diachronic reading of the EXPRESSIVIST THESIS that we all *now* use moral language to a-express *beliefs* of ours about things. Initially, this might seem like exactly the sort of outcome the expressivist should want. After all, his semantic theory is supposed to be one that is compatible with the possibility of lying (or otherwise linguistically misrepresenting one's mental states). And this is only possible if we separate the meaning of a moral claim from whatever mental states are occurrent in a speaker's mind as he now makes it. The problem, though, is that the presence of the relevant mental states in competent users of moral language *has* to be more important to the expressivist than this suggests. First of all, if expressivism is really just a theory about the sort of semantic contents that moral claims s-express, and those contents are supposed to be provided entirely by the genetic heritage of the claims and

conventions governing their use, then why should the expressivist be bothered at all by users of

the language who do not now have the relevant mental states to a-express? As an example,

Blackburn imagines a case that is supposed to be analogously relevant to (and problematic for)

an expressivist treatment of aesthetic language:

> Magnus, let us suppose, always had a passion for wine. Furthermore, he
> discovered early in life that he had a very good memory for taste. So in due
> course he became a wine connoisseur, a Master of Wine, and exercised his taste
> and skill to build up business as a writer, shipper, and expert on all matters
> related to wine. So far so good. But now we imagine a gradual change in
> Magnus. He finds that his memory and powers of discrimination are as strong as
> ever. But he finds his enjoyment slipping away. The years of slurping and
> spitting have taken their toll. He now goes to tastings with a heavy heart, when
> before he would have enjoyed them greatly, and this although he recognizes that
> the average quality of wine on offer has gone up remarkably in his lifetime. He
> takes water with his meals, and drinks only scotch before and after.[8]

But, you might wonder, why should this be a problem? If aesthetic expressivism amounts to

nothing more than the view that aesthetic claims s-express contents that relate only to the

conative attitudes of those who first used aesthetic language (and not any of Magnus' mental

states), then to *what*, exactly, does Magnus threaten to be a counterexample? Nonetheless,

Blackburn writes,

> [a]n aesthetic expressivism may have problems with Magnus. Let us suppose
> such an expressivism built around the plausible idea that to say that a wine is
> good is to express or voice pleasure in its taste, and to say one wine is better than
> another is to externalize, that is, to voice or communicate more pleasure in its
> taste, and so on. Such a theory, obviously, has the same powerful motivation
> behind it as ethical expressivism, namely that it puts the pleasure we get from
> wine as fundamental to the exercise of grading it. And that seems right. … But
> now it seems to follow that Magnus can't mean what he says if he pronounces a
> wine good, or better than another, because the requisite pleasures have
> disappeared from his life.[9]

This just doesn't make any sense, though, if expressivism really is supposed to be understood in

terms of the diachronic story I outline above. After all, according to that story, when Magnus

---

[8] Blackburn unpublished: 8.

[9] Blackburn unpublished: 8.

17

claims of one wine that it is good, *he is expressing approval of it*, where this means that his claim s-expresses a content that has been provided for him by the approving attitudes of those who first used the language. Magnus' own attitude, or rather lack thereof, is totally irrelevant. Magnus' mental states—and by analogy, the mental states of speakers who now use moral language—should only seem relevant to expressivists (be they aesthetic or ethical expressivists) if expressivism involves more than just the conative attitudes of the original speakers of the language.

Relatedly, if the expressivist is to make sense of the alleged tie between expressivism and motivational internalism, then there will have to be more that just a passing concern on his part that users of moral language do so *with* the relevant attitudes to a-express. In other words, if moral language does not function (at least partly) to express *my* conative attitudes (and not just the fossilized attitudes of those who first used it), then why should we expect my use of moral language to be accompanied by any motivation on my part to behave in accordance with the normative force of the claims I make? One of the virtues of expressivism was supposed to be that it could account easily for internalism, but in order for it to do so, it ought to be the case that moral claims function to express more than just the shadows of the attitudes of those who first made them. This all suggests that the EXPRESSIVIST THESIS ought to be understood along different lines than the ones suggested by the diachronic story. So how else might the expressivist go about answering the expression question?

One way to ensure that speakers' attitudes play a more significant role in our understanding of the EXPRESSIVIST THESIS is to tell a more *synchronic* story about how moral language functions to express the conative attitudes of those who now use it. And there are different ways of doing this. One implausible synchronic story about the expressive function of moral language might look something like this: the semantic content that a moral claim now s-

expresses depends upon whatever mental state I am using it to a-express. So if, in claiming that murder is wrong, I am a-expressing a negative attitude toward murder, then my claim is here synonymous with something like "Boo to murder!" But if, in claiming that murder is wrong, I am a-expressing a positive attitude toward Gorgonzola cheese, then, in this instance, the claim would be synonymous with something like "Hooray for Gorgonzola cheese!" This account, unlike the diachronic account described above, would give speakers' mental states a more prominent role to play in the expressive function of moral language, but it's also problematic for a number of obvious reasons. For starters, it still doesn't do enough to explain the supposed tie between expressivism and internalism. Expressivists surely cannot lay claim to the sort of internalism they want if moral language can properly function to express any ol' state of mind. If speakers who use the sentence "Murder is wrong" to a-express a negative attitude toward murder are using the language *no more correctly* than those who use it to a-express a positive attitude toward Gorgonzola cheese, then we haven't got any more reason to expect speakers who utter the sentence "Murder is wrong" to me motivated to avoid murdering than we have for expecting them to eat Gorgonzola cheese. It is not enough that moral language functions to express the conative attitudes of those who *now* use it if their use of it is not governed by any conventions. Furthermore, this synchronic story may not be compatible with the possibility of lying (or otherwise linguistically misrepresenting one's mental states). For if the semantic contents of the claims I make depend entirely upon whichever of my mental states I use those claims to a-express, then it is hard to see how any claim of mine could possibly mean something that is incongruous with my mental states. For these and other reasons, I suspect this will not be the sort of synchronic story that the expressivist will want to tell in response to the expression question.

There is another, and I think much more plausible, synchronic story for the expressivist to tell. But before I explain what that story is, it will help for me to distinguish between the *honest* use of expressive language and the *sincere* use of expressive language. Consider a person who, as a kind of reflex, winces or cries out before the dentist has even touched his teeth and gums with those pointy instruments of pain. And compare that to a case in which someone winces or cries out in order to trick an onlooker into thinking that he's in pain. For my purposes, I want to say that (a) both of these people are being insincere, since neither of them actually has any pain to a-express, but (b) only the latter of the two is being dishonest. A speaker uses expressive language dishonestly if he does so with an intention to deceive (or otherwise misrepresent himself), while a speaker uses expressive language insincerely if he does not have the relevant mental state to a-express. So with this distinction in hand, here is the second and more plausible synchronic story: the semantic content that a moral claim now s-expresses depends upon whatever mental state has been assigned, by convention, to the sincere production of that claim, so that a speaker (who is using the language honestly) does not use the language *correctly* unless she uses it to a-express the relevant mental state. So, for instance, if convention has it that the sentence "Murder is wrong" is now synonymous with something like "Boo to murder!" then, as long as I am being honest, either (a) I use that sentence to a-express my disapproval of murder, or (b) I am misusing the language. This sort of view has a number of advantages over the diachronic and synchronic accounts described above. For one thing, it sets aside as secondary the issue of how these conventions came about. The expressivist is of course free to tell the sorts of just-so stories that Blackburn and Gibbard have told, but nothing really hinges on this. All that really matters is that there are reasons for thinking that it now functions to express conative attitudes. This synchronic account also makes sense of why cases like that of Magnus the disaffected wine critic are a problem for the expressivist. If Magnus is not using the

20

sentence "This wine is good" to a-express his pleasure, then he is either being dishonest or he is misusing the language. But, you might think, Magnus is neither being dishonest nor is he misspeaking in any way. As Blackburn himself allows, "his memory and powers of discrimination are as strong as ever," and so "Magnus is, surely, still perfectly entitled to his place as a wine critic."[10] This synchronic account also explains well the expressivist's claim to internalism. If, when speakers (correctly) make moral claims, they a-express corresponding conative mental states, then that would account for the alleged conceptual tie between moral claims and motivation, since conative mental states are generally thought of as motivational states. And finally, this synchronic account is perfectly compatible with the possibility of lying, i.e., using the language dishonestly.

So I think this latter of the two synchronic stories is the one that expressivists ought to tell in their response to the expression question. And so, the expressivist's answer to the expression question looks something like this: when the EXPRESSIVIST THESIS says that the function of moral language is to express only conative attitudes, that should be taken to mean that moral language functions, first and foremost, to allow speakers to a-express only their conative attitudes toward things, but it does so in a way that is governed by convention, so that we either regularly use moral language to a-express the relevant conative attitudes or we regularly misuse the language. Now, I think we can set aside the latter of these two possibilities. It may be the case that we are all misusing moral language, but I'll assume for now that expressivists would rather not commit us all to any sort of systematic linguistic incompetence. So it comes to this: according to the EXPRESSIVIST THESIS, whenever a speaker uses moral language honestly, she a-expresses only conative attitudes. In the next two sections, though, I'll take a much closer look at this notion of a-expression. And it will turn out in the end, I think, that a satisfactory

---

[10] Blackburn unpublished: 8.

account of what is really involved in a-expression actually spells serious trouble for the expressivist.

4. WHAT IT TAKES TO A-EXPRESS

In order to better understand the notion of a-expression, i.e., what it takes for speakers to express mental states, there are three questions that I will ask in this section. First, can speakers express mental states that they do not in fact have? Second, does their ability to express a mental state depend upon their behaving in a manner recognizable to others as an expression of that state? And third, what is the connection between expression and speakers' intentions? (For the record, from this point forward, I'll only be speaking of "expression" in the sense of a-expression, unless otherwise noted.)

It seems to be a rather common assumption among philosophers of language, and especially those with interests in the nature of expression, that speakers can indeed express mental states that they do not have. Maybe the assumption goes back to Searle, who wrote in his book *Speech Acts*,

> Wherever there is a psychological state specified in the sincerity condition [of a speech act], the performance of the act counts as an *expression* of that psychological state. This law holds whether the act is sincere or insincere, that is whether the speaker actually has the specified psychological state or not. Thus to assert, affirm, state (that *p*) counts as an expression of belief (that *p*).[1]

And Blackburn echoes,

> Few philosophers find it problematic that a sentence can express a belief, although the person uttering it not believe what he thereby expresses. Insincere expression is expression nonetheless. And it is the same with other things we do with words. A warm greeting expresses warmth, even if the greeter feels none.[2]

I don't think this is right. Consider the following three cases:

---

[1] Searle 1969: 65.

[2] Blackburn unpublished: 8.

I.   Tom experiences a pain and consequently says, "Ouch!"

II.   Jillian does not experience a pain, but still says, "Ouch!"

III.   Steve experiences a pain and consequently says, "Kelly Clarkson!"

As I see it, the speech acts in cases I and II have something important in common that is absent in case III. And the speech acts in cases I and III have something important in common that is absent in case II. Now, according to Searle, since the experience of pain constitutes the sincerity condition for the speech acts in cases I and II, both acts count as expressions of pain—the first one sincere, the second insincere. But presumably, since there just aren't any conventions of sincerity that associate the saying of "Kelly Clarkson!" with the experience of pain, it is either the case that (a) Steve isn't expressing pain or (b) if Steve's speech act *does* count as an expression of pain, it is in virtue of something other than sincerity conditions. So the distinction between sincere and insincere expression is perhaps helpful for comparing and contrasting cases I and II, but it offers nothing by way of explaining what is going on in case III, and, importantly, how it relates to the other two cases. If, however, we distinguish between a speaker's *expression* of a mental state and what we might call a speaker's *acting-as-if* she has a particular mental state, then I think we will be able to tell a much more complete story about the relative similarities and differences at work among the three cases. Suppose that a speaker acts-as-if she has a particular mental state just in case she performs the sort of action, linguistic or non-linguistic, that, in ordinary circumstances, an audience would likely interpret as an indication that she has that mental state (perhaps because of various conventions, social expectations, etc., that associate that sort of behavior with that particular mental state). So Tom and Jillian are acting-as-if they are in pain, but Steve is not. And this accounts for what cases I and II have in common, but case III lacks. But this is not to say that Steve is not expressing pain. After all, the saying of "Kelly Clarkson!" is, from his perspective, just as much an expression of his pain as "Ouch!" is

for Tom. It is, if you like, Steve's own idiosyncratic way of *airing* or *giving vent to* the pain that he experiences.

Some, however, may want to deny that Steve expresses pain. Mitchell Green, for instance, writes,

> [T]o express my psychological state I must make it knowable to an appropriate observer. Although the relation between expression and what is expressed is not always so straightforward as the relation between tears and grief, I suggest that one theme binding together different forms of expression is the ability of the expressive behavior or artifact to convey, or at least enable in an appropriate observer, knowledge of what it expresses. … I express a psychological state, then, by showing it in such a way as to enable propositional knowledge of it in appropriate observers.[3]

But what about idiosyncratic behaviors, like Steve saying, "Kelly Clarkson!"? Here is what Green says about idiosyncratic expression:

> Although an important class of expressive behaviors are pan-cultural, we also express ourselves in ways that are either idiosyncratic to an individual, family, or the like, or subject to conventions local to our culture or ethnicity. … In all these cases [of idiosyncratic expression] we can nevertheless show [i.e., express] what is within, *albeit only in ways intelligible to those familiar with our idiosyncrasies or conventions.*[4]

So Steve's ability to express his pain, according to Green, depends upon whether or not Steve's action is recognizable as an expression of pain, idiosyncratic as it is. And this depends upon there being at least some kind of convention or mutual understanding in place between Steve and whoever it is that is familiar with his idiosyncrasies. For otherwise, Steve is not making his pain knowable, and so, he is not expressing it. Green, then, seems to be suggesting that one cannot express a mental state unless one acts-as-if one has the mental state, since we only make our mental states knowable to an audience if we perform the sort of action that an audience would likely recognize as an indication that we have that mental state. But why think a thing like this?

---

[3] Green 2007a: 270.

[4] Green 2007b: 19, italics mine.

Here is my guess. I think philosophers assume that expression and what I'm calling acting-as-if must be somehow or other tied up together because they see expression as an essentially *communicative* endeavor. Steve cannot express pain without acting-as-if he's in pain because, well, how else could he possibly communicate to an observer that he is in pain? The problem with this, though, is that there seem to be cases in which expression and communication could come apart. When I stub my toe, for instance, there's really no telling what sorts of strange noises might come out of my mouth. Maybe, when I say, "Ouch!" I express pain. But what if I say something similar, but not identical, to "Ouch!"? And what if I say something slightly less similar than that? Are we to say that there is a point at which I would no longer be expressing my pain, because, at that point, my speech act would no longer be the sort of thing that would (or could) communicate to an observer that I am in pain? That seems odd. It might be the case that my ability to *communicate* that I am in pain depends upon my action being sufficiently in line with conventions, or expectations, or whatever; but why suppose that I can only use *certain* words to express my pain, and not others? Suppose I shout a series of exclamations, beginning with "Ouch!" Then, "Oush!" Then, "Oosh!" And finally, "Oolsh!" Do I somehow go from *expressing* to *not expressing* pain, because shouting "Ouch!" is a way of acting-as-if I am in pain but "Oolsh!" is not? Again, I might, in some sense, go from *communicating* that I am in pain to *not communicating* that I am in pain, and this might in fact be the result of my having strayed too far away from convention. But it seems really strange to suppose that my ability to *give vent to*, or *press out*, or *release* my pain similarly depends upon me saying what people *conventionally* say when they are in pain, or what people would *expect* me to say when I am in pain, etc.

It is one thing to suppose that, oftentimes (or even, *always*), when I express a mental state, I make it knowable to some appropriate observer that I have that mental state. It is quite

another thing, though, to suppose that I *cannot* express a mental state *unless* I make it so knowable. Imagine a happy man, happily walking through a park, whistling a happy tune. Is he expressing his happiness? I think he is. Is he making his happiness knowable? I suppose so. But is he only expressing his happiness *because* he's making it knowable, i.e., is the fact that he's making his happiness knowable one of the things *in virtue of which* he counts as expressing? I don't think so. And the only reasons I can think of for supposing that to be the case all involve the assumption that expression is an essentially communicative endeavor. Here is another case to think about.[5] Imagine someone whose native language is something other than English trying to learn various English idioms. She's heard people say, for instance, "Well, I'll be darned" a few times, but not clearly enough to make out the actual words. And we'll suppose that she's got a clear enough idea of why people typically use the phrase—to express their surprise. So, feeling surprised, she tries to put that idiom to use, but since she hasn't yet discerned the actual words, she says something that is only phonetically similar to it: instead of saying, "Well, I'll be darned," she says, "Well, ah be dah." Is she expressing her surprise? Again, I think she is. But is she making her surprise knowable? It's hard to say. So in virtue of what, exactly, is she expressing surprise? In conversation, Green has suggested that this does indeed count as an expression of surprise in virtue of there being a kind of causal-explanatory connection between her utterance and the actual idiom, i.e., were there *not* this causal-explanatory connection, she would not be expressing her surprise because she would not be making her surprise knowable to an appropriate observer. This all seems really strange to me. It might be right to say that she makes her surprise knowable because of the causal history of her utterance, but to suggest that this causal history is the thing in virtue of which she *expresses* her surprise seems to me to rest way too much of our ability to express our mental states upon our ability to communicate those

---

[5] I owe this one to Dorit Bar-On.

states to others. How odd it would be to say something like, "It's a good thing her utterance bears this interesting causal relationship to the phrase 'Well, I'll be darned'. For otherwise, she wouldn't have been expressing her surprise." Whether or not she's expressing her surprise ought to be more *up to her* than such a suggestion allows.

Let me make one further point. Oftentimes, when people write about expression, they use various other words or phrases that are supposed to be synonymous with "express." I've already used *air*, *give vent to*, *press out*, and there are many others: *convey*, *reveal*, *evince*, *show*, *put out there*, *present*, etc. It's interesting to note that some of these terms seem to have packed into them already an assumption about there being some relationship in place between a speaker and an audience. *Showing*, *presenting*, *revealing*, etc., are typically thought of as things we do *to* an audience. If I am showing a picture of my family, presumably there is someone *to whom* I am showing the picture. But there is not this sort of audience-involved assumption packed into some of the other terms. Presumably, I can *give vent to* some internal state without there being someone *to whom* I am venting that state. So when, for instance, Green constructs an account of expression that is based heavily upon the notion of showing,[6] he may backhandedly be assuming that we cannot *express* a mental state unless there is someone *to whom* we could communicate that we have that state. Expression may be related to communication in various interesting ways, but, as I've argued above, there seem to be cases in which the two come apart. So the relationship between expression and communication cannot be as close as Green and others think it is. Again, it may be that, in every case in which I express a state of mine, there is the possibility for communication. But this is not to say that my ability to express depends upon there being the possibility for communication. And this leads me to think that my expressing a mental state does not, in fact, depend upon my acting in an appropriately conventional sort of way, or in

---

[6] See, for instance, Green 2007b: 18, 22, 43, 46-53.

accordance with certain social expectations, etc. Expressing and acting-as-if are just two distinct types of activity.

So, to return to the three cases that we were considering above: maybe, by acting-as-if she is in pain, Jillian communicates that she is in pain. And maybe, since he does *not* act-as-if he is in pain, Steve fails to communicate that he is in pain. But this is not to say anything about who is expressing pain and who is not. So, until I am given some other reason for thinking that acting-as-if is either necessary or sufficient for expression, I'll assume that it is neither. And now I've got an easy way to tell a more complete story about the relative similarities and differences at work among cases I, II, and III. Tom and Jillian are *acting-as-if* they are in pain, but Steve is not. And Tom and Steve are *expressing* pain (we'll assume for now), but Jillian is not.[7] So, in order for a speaker to a-express a mental state by performing some speech act, the act must be sincere (contra Searle), but it need not be recognizable to others as an expression of that state (contra Green).

You might be asking why this is important. Green, for instance, thinks the issue of whether or not insincere speech acts ought to count as legitimate acts of expression is "largely a matter of nomenclature,"[8] but I think that's probably an understatement. There's more going on here than just a quibble about what we should *call* insincere speech acts. If it turns out, as I'm suggesting, that (a) a speaker can only a-express a mental state if the speaker *has* the relevant mental state to express, and that (b) acting in accordance with the conventions of one's language, or in accordance with the expectations of one's linguistic community, etc., is neither necessary

---

[7] Some will say that, while Jillian may not be *expressing* pain, her utterance is nonetheless *expressive* of pain (see, for instance, n. 41). I will avoid the notion of expressiveness, however, primarily because I want to be sure not to confuse expressiveness with expression. The two are distinct and I am here only interested in the latter.

[8] Green 2007a: 281, n. 8. Here Green writes, "As we have seen, Searle says that an assertion, for instance, is an expression of belief even when the speaker lies. How we treat such a case is largely a matter of nomenclature: I prefer to say that in such a case the assertion is *expressive* of belief but does not express the speaker's belief (for he has none)."

nor sufficient for a-expression, then it begins to seem as though a-expression is something that depends more upon what is going on in a speaker's head than upon anything outside of the speaker. In other words, it is beginning to seem as if whether or not I express a particular mental state by performing some speech act is, in some important sense (that I will flesh out in the next section), *up to me.* And this is significant because, quite frankly, if what I understand to be going on in my head when I use moral language is something other than what the expressivist suggests is going on, then either (a) I'm somehow or other systematically confused about my own mental states, which would be a troubling discovery indeed, or (b) expressivism is false. Or so I will argue in the next section.

There is, however, still the issue of intention. To what extent should we understand the expression of a mental state, or the performance of a speech act that expresses a mental state, to be an intentional phenomenon? By most accounts, a speaker need not intend to express a particular mental state in order for that mental state to be expressed. But that is not to say that the sense of expression with which we are here concerned is wholly independent of the intentions of speakers. What, then, *is* the relationship between a-expression and speakers' intentions? It may help, at this point, for me to rehearse a somewhat familiar semantic distinction in order to get our hands on a proper understanding of the relationship between expression and speakers' intentions. In his 1957 paper "Meaning," H. P. Grice distinguishes between two senses of the word 'mean' (and its cognates).[9] There is first the sense in which, for instance, we say,

3.   "Those spots *mean* measles,"

or,

4.   "The recent budget *means* that we shall have a hard year."

---

[9] Grice 1957: 377-378.

According to this sense of 'mean', the spots and budget serve as signs or indicators of measles and an impending hard year respectively, but their doing so is not the result of anyone's intentions. Grice calls this *natural meaning*. Alternatively, there is a different sense of 'mean'—one that Grice calls *non-natural meaning*—at work in such statements as

5. "Those three rings on the bell (of the bus) *mean* that the 'bus is full',"

or,

6. "In gesticulating that way, Salvatore *means* that there's quicksand over there."[10]

Non-natural meaning is somehow or other a matter of someone's intentions, e.g., the intentions of the bus driver in ringing the bell (or maybe the intentions of designer of the bell-ringing system), the intentions of Salvatore in so gesticulating, etc. Grice himself actually suggests that non-natural meaning involves a rather complicated set of intentions on the part of a speaker, and many have since wondered if Grice's own construal of the sorts of intentions necessary for non-natural meaning is ideal.[11] For my part, though, all that matters is that, first, expressions of mental states can form subsets of cases of both natural and non-natural meaning; and second, the sort of expression relevant to my study seems only to be a case of non-natural meaning.

Upon having his pants pulled down by a prankster in a very public setting, William first blushes and then says, "That pisses me off." William's blushing expresses his embarrassment, and then, by saying what he says, he expresses his anger. But the blushing is more a *natural* sign of, or manifestation of, embarrassment (like the spots are of measles), while William expresses his anger in a *non-natural* sort of way. In other words, if William's blushing means that he is embarrassed, it does so in the natural sense of meaning; and if William's utterance means that he is angry, it does so in the non-natural sense of meaning. So expressions of mental states can fit

---

[10] This last example is taken from Green 2007b: 54.

[11] See, for instance, Bar-On 1995 and Green 2007b: 53-81.

either natural or a non-natural descriptions. But the sort of expression relevant lying at the heart of the EXPRESSIVIST THESIS, I submit, will always and only be non-natural. The EXPRESSIVIST THESIS is supposed to apply only to speech acts that are every bit as intentional as William's angry utterance. We can, of course, imagine someone unintentionally using moral language, as in a Tourettic outburst or something. But in such a case, the speaker would not be expressing any of her conative states. She may be *acting-as-if* she has some conative state, and she may utter something that s-expresses a content that is somehow or other related to a particular conative state; but she is not *expressing* any state of hers.

So in order for moral language to fulfill its role of allowing speakers to express their conative states, the speaker must both (a) *have* the attitude to be expressed and (b) put the language to use in the *intentional* performance of a speech act. But this is still not enough. Suppose my friend Liesl is visiting me from her home country of Austria, and she knows very little English. Suppose also that she's strongly anti-abortion. One evening, I'm hosting a dinner party so that some of my American friends can meet her, and at one point, Liesl asks me how to say a pleasant English greeting. Feeling mischievous, I respond, "Try saying, 'Abortion is evil'." So Liesl practices a few times, turns to one of the American partygoers, and says, "Abortion is evil." Here is a case in which a speaker both (a) has a negative attitude toward abortion and (b) puts the term 'evil' to use in the intentional performance of a speech act, and yet, it does not seem right to say that Liesl is here expressing her negative attitude toward abortion. And the reason for this, it seems, has something to do with the fact that Liesl doesn't really know what she's doing. This suggests that a further condition must be added to my analysis of a-expression, one that places the speaker in the right sort of epistemic relationship to her own speech acts and the mental states she is supposed to be expressing by performing those acts. In the next section,

I'll offer and defend a proposal for how this relationship ought to be understood; and then I'll

argue that the proposal spells serious trouble for expressivism.

5. A-EXPRESSION AND KNOWING ONE'S REASONS

Throughout the past few decades, philosophers and cognitive scientists alike have been paying more and more attention to the epistemology of action and the epistemology of language. With respect to action, for instance, there are studies of an agent's awareness of his own intentions, the sense an agent has of control, the role and nature of perception in action, and various other phenomena. And with respect to language, there are studies of what it takes to know a natural language, what it takes to know the meaning of language, and the dependence of various mental states like beliefs and desires upon an agent's knowledge of language and meaning. And of course, there is also the much-discussed matter of self-knowledge, i.e., knowledge of one's own mental states. For better or worse, my interest here—i.e., a speaker's expression of a mental state via some intentional speech act—appears to fall somewhere in the middle of all of this. But rather than rehearsing various issues that have sprung up in debates about self-knowledge or the epistemologies of action and language, what I'll do in this section is the following. First, I'll put forth my proposal for the sort of relationship that a speaker must stand in with respect to her speech acts and her mental states in order for her to express those mental states by performing those speech acts. My hope is that this will require only the most superficial of engagements with some of the literature I've just mentioned, since a full immersion in the material would likely throw this discussion way off course. Second, I'll defend my proposal by showing how it answers to our intuitions in some of the cases I've already described and even a few others. And then finally, I'll explain how the proposal amounts to a serious problem for expressivism.

As you recall, by the end of section 3, I concluded that the expression of a mental state, in the relevant sense of expression, involved at least these two things: (a) the possession of the mental state and (b) the intentional performance of some speech act. But this was not yet a complete picture of a-expression. In order for a-expression to take place, it must also be the case that the speaker stand in some sort of relationship to both the speech act and the mental state that is to be expressed. Here is how I propose we understand that relationship:

A speaker S a-expresses some mental state M by performing speech act φ iff
    (i) S has M,
    (ii) S *owns* her φing, [which S does iff
        (a) S φies intentionally, and
        (b) S knows (even if only tacitly) her reason(s) for φing], and
    (iii) M is among S's reasons for φing.[1]

Now let me unpack this a bit. At work in condition (ii) is a concept that philosophers of action have put to use for various purposes, namely, that of the *ownership* of action. Roughly speaking, to say that a person owns an action is to say that the person is aware of the action as one's own. But the matter is not nearly as simple as this may sound. Some, for instance, make a distinction between the *fact* of ownership and the *sense* of ownership; others make a distinction between a person's awareness of *what she is doing* and her awareness *that she is the one doing it*; and there are still other distinctions that get bandied about in discussions of agency and ownership. What seems to lie just below the surface of most conceptions of ownership, however, is an idea that goes something like this: in order for an agent to own one of her actions, not only must she have acted intentionally, but she also must be aware of her reasons for so acting (even if she is not consciously aware of those reasons in the moments leading up to her action).

---

[1] Two quick notes. First, "φing" should not here be understood as standing in the place of some speech act *type* (e.g., asserting), but rather, some speech act *token* (e.g., Joe's asserting that it is raining outside). And second, I take this account of expression to apply just as well to many non-linguistic acts as well. My intentionally smiling expresses happiness (or some such conative state) if (a) I *am* happy, (b) I *own* my smiling, and (c) my happiness is among my reasons for smiling.

On [Lucy] O'Brien's view, you know what you are doing in virtue of doing it with a sense of control, where this means that you act on the basis of an assessment of the options—that is, on the basis of an answer to a practical question. In [Johannes] Roessler's account, a crucial part is played by agents' use of perceptual evidence to answer practical questions. … Given that answering the deliberative question seems to involve evaluating your practical reasons, there is … an essential normative ingredient in the sense of ownership. Crudely, to be aware of an action as one's own is to be aware of it as controlled by one's own practical reasons.[2]

Imagine, for instance, two people making the thumbs-up gesture. Trevor gives his friend a hearty thumbs-up in a show of congratulations, but Anna has no idea why she's making the gesture. She's suffering from anarchic hand syndrome. Setting aside for the moment the issue of intention,[3] it seems clear that if either of these actions has any chance of being owned, it is Trevor's, since he is the only one of the two thumb-raisers that is in a position to articulate his reason(s) for doing so. And we can see immediately how this is relevant to speech acts and a-expression. Recall, for instance, the example of the person shouting, "Murder is wrong!" in a Tourettic outburst. We said that this person, unlike the one who says, "Murder is wrong" in normal circumstances, does not in fact express her disapproval of murder. Similarly, Anna, unlike Trevor, does not express a congratu-latory attitude when she raises her thumb. And this is because, even if they actually *have* the relevant states of mind (e.g., imagine Anna actually having a desire to congratulate her friend at the very moment that her anarchic hand syndrome kicks in), neither Anna nor the person with Tourettes is in a position to justify either action by appeal to reasons. One's ownership of an action, we might say, entails one's having the resources to *rationalize* the action.

We can return now to the cases I, II, and III from above. Earlier, we decided that Tom and Jillian, by saying "Ouch!" both act-as-if they are in pain, but Steve does not. And Tom and

---

[2] Eilan and Roessler 2003: 45, 46. Here the authors refer to O'Brien 2003 and Roessler 2003.

[3] It may seem clear that Anna's thumb-raising is unintentional, and I think that is probably the consensus view, but there is also the complicated matter of unconscious intentionality (which I will be avoiding here).

Steve both express pain, but Jillian does not. Now we can see that this is because Tom (we can assume) knows that his experience of pain is his reason for saying "Ouch!" And the same goes for Steve: he knows that his pain is his reason for saying "Kelly Clarkson!" Jillian, though, could not know such a thing, since she is not experiencing any pain. (Of course, there may be some other mental state that is serving as her reason for saying "Ouch!" Perhaps she is an actress, and it is her desire for the audience to believe that she is in pain that is her reason for saying "Ouch!" And so, maybe she expresses a desire of this sort. But she does not express pain.)

Consider, again, the case of Liesl described at the end of section 3. Recall that she has a negative attitude toward abortion, and I've tricked her into saying, "Abortion is evil" to various partygoers. But we do not want to say that Liesl is here a-expressing her negative attitude toward abortion. And this is because, even though she's got the attitude to be expressed, she will not recognize that attitude as one of her reasons for saying what she's saying to the partygoers. I may only be pulling this prank because I know that she has a negative attitude toward abortion (and perhaps I know that all of my other guests are vehemently pro-choice); and so, there may be a sense in which her attitude is a kind of distal explanatory reason for her actions. But in order for Liesl to be expressing her attitude toward abortion, she herself has to know (even if only tacitly) that she's saying, "Abortion is evil" *because* of her attitude toward abortion. And of course, in this situation, she can't know such a thing, since it isn't the case. So not only does my analysis of a-expression appear to account quite smoothly for ordinary cases like Tom's expressing pain by saying "Ouch!" but it also seems fit to handle all sorts of non-ordinary cases, including (a) insincere speech acts (e.g., Jillian's saying "Ouch!"), (b) idiosyncratic expression (e.g., Steve's saying "Kelly Clarkson!"), and (c) cases in which there is some sort of mismatch between a speaker's speech act and the mental state that the speaker has before her

37

own mind as she acts (e.g., Liesl's thinking that she is greeting the partygoers by saying to them, "Abortion is evil").

How does this analysis of expression constitute a problem for expressivists? In section 2, we decided that the best way for the expressivist to account for the expressive function of moral language was in terms of a synchronic story according to which the semantic content that a moral claim now s-expresses depends upon whatever mental state has been assigned, by convention, to the sincere production of that claim, so that speakers who use the language honestly a-express only the relevant conative attitudes. Here is the problem, though: if, on any occasion in which I put moral language to use in the intentional performance of a speech act, I recognize some *belief* of mine as being among my reasons for performing the speech act, then either (a) I am, on such an occasion, not only unable to identify which of my own mental states is being expressed, but also unable to recognize my own reasons for performing that intentional action; or (b) the EXPRESSIVIST THESIS is false, since, on such an occasion, moral language would here be functioning to express a belief. Of course, the expressivist might be right about some cases of speech acts involving moral language. Maybe, in some cases, when I say that such-and-such is morally wrong, my reason for doing so is my conative attitude toward such-and-such. But if there is ever a case in which some *belief* of mine about such-and-such is among my reasons for saying that it is morally wrong, then, again, the EXPRESSIVIST THESIS would be falsified, since I would here be using moral language to express a mental state that is not conative. To borrow Blackburn's words, the EXPRESSIVIST THESIS is the expressivist's attempt to "offer an explanation of what we are up to in going in for this [moral] discourse."[4] Surely, though, what it is that I am up to on any given occasion when I intentionally use moral language

---

[4] Blackburn 2007.

ought to be, to some degree at least, *up to me.* Generally speaking, I'm the authority when it comes to my own mental states, intentions, and reasons for acting. Or so we might think.

Consider another case, similar to the case of Liesl and her negative attitude toward abortion. My friend Sean and I are sitting in a café in France, but I am the only one of the two of us who knows any French. There's a woman at a nearby table that Sean finds attractive, and so he turns to me and asks, "How do I tell that woman that I think she's beautiful?" Unfortunately for Sean, though, I seize the opportunity for mischief and respond, "You just point at her face and say, 'C'est dégoutant, ça! Berk!'" So when he feels that the time is right, Sean turns to her, points at her face, and says, as smoothly as possible, "C'est dégoutant, ça! Berk!" But to his surprise, the woman reacts in horror. So he turns to me and asks, "What did I just do?" I might answer, "The sentence 'C'est dégoutant, ça! Berk!' really means 'That's disgusting! Yuck!' So even though you thought you were expressing a *positive* attitude toward her appearance, you were actually expressing a *negative* attitude toward her appearance." Something's amiss here, and I think Sean would be right to protest. After all, he takes himself to be an authority when it comes to the states of his own mind that get a-expressed when he performs intentional actions. So he might insist, "I don't even have a negative attitude toward her appearance. In fact, my only reason for speaking to the woman in the first place was the positive attitude that I have toward her appearance. So I don't care what the words mean, or what sort of attitude people usually express when they intentionally use the words I used. I was *not* expressing a negative attitude. I may have communicated to her that I have such an attitude, but that's not the sort of attitude that I was expressing." Of course, it is *not* up to Sean whether or not the thing he uttered s-expresses a content to which convention has assigned the attitude of disgust. But it *is* up to Sean whether or not he is a-expressing such an attitude.

Similarly, in response to the expressivist, there may be occasions for speakers to make such an appeal to their own first-person authority. Suppose I have a negative attitude toward the mass breeding of dogs in puppy mills. And suppose, in conversation, I intentionally say, "The mass breeding of dogs is morally wrong." If, on this occasion, I know my reasons for saying what I say, but my distaste for mass breeding is not among them, then I do not here a-express my negative attitude toward the mass breeding of dogs. Imagine, for instance, that the context of the conversation is not of the sort in which my abhorrence of mass breeding would be anywhere near the forefront of my mind. Maybe I'm just calmly rattling off a random list of things I believe to be morally wrong (if you'd like a mental image: picture Bubba rattling off a list of dishes containing shrimp in the movie *Forrest Gump*). Of course, there may be other conative attitudes of mine that are among my reasons for saying what I say, like, e.g., a desire to share my moral beliefs with others. But it does not seem right to say that my abhorrence of mass breeding is, on this occasion, among my reasons for saying what I say. Nor would it seem right to say that I am being either dishonest or insincere. I very much mean what I say.

Now, of course, the expressivist could insist that, whether or not I *recognize* my abhorrence of mass breeding as one of my reasons for saying what I say, it *is*, in fact, one of my reasons for doing so. But I think this is doubly problematic. First, it straightforwardly conflicts with the idea that speakers have a kind of first-person authority when it comes to such things as their own mental states and their own reasons for acting. And second, as a consequence, it commits speakers to a rather grievous error. Even if I sincerely believe that I'm saying, "The mass breeding of dogs is morally wrong" because of my belief about the wrongness of mass breeding, the expressivist is suggesting that I'm just wrong about that. Unbeknownst to me, I'm *really* saying what I say because of my abhorrence of mass breeding. This is significant because, as you will recall from my introduction, expressivism was motivated in the first place partly by a

desire to avoid committing competent users of moral language to the sort of systematic error to which Mackie had us all committed. According to Mackie, all of our moral thought and discourse rests on a mistake, since it fails to actually refer to any real facts or properties in the world. And expressivism was supposed to avoid this conclusion by denying that moral language was ever meant to refer to real facts or properties in the world. But now we see that expressivists have only saved us from one error by committing us to another. Now, whenever we take ourselves to be using moral language because of various beliefs of ours (just as we do when we talk of, say, the shapes of things), we're just *wrong* about that. And arguably, this sort of mistake is even more disturbing than the one to which Mackie had us committed, since this is a mistake about things we ordinarily take ourselves to be aware of via simple introspection. It is one thing to use language on the mistaken assumption that it refers to things in the world—if the opinions of some are right, theists make this mistake on a fairly regular basis. It is another thing, though, to be mistaken about things over which we ordinarily take ourselves to have a significant measure of authority—namely, our mental states, our intentional actions, and our reasons for acting. So either (a) competent users of moral language are often in error with respect to their own mental states, their own intentional speech acts, and their own reasons for performing those speech acts, or (b) speakers can sometimes express beliefs when they use moral language, in which case the EXPRESSIVIST THESIS is false.

6. SOME WAYS FOR THE EXPRESSIVIST TO RESPOND

In this section, I want to consider a few ways in which the expressivist might respond to the case I've been making against expressivism. Since most of the heavy lifting is actually done by my analysis of the nature of a-expression, the expressivist may want to rebut the analysis. And of course, this might be done in any number of ways. For now, though, I'll assume that each of the three parts of my analysis has been adequately defended: a speaker cannot express a mental state that she does not have; a speaker must own a speech act in order to express a mental state by performing the act; and the mental state to be expressed must be among the speaker's reason for performing the act. There still, however, may be other causes for concern. For instance, according to some, a satisfactory analysis of expression ought to allow for cases in which a person intends to express some mental state but fails to do so. This is one of the issues that Green raises in his criticism of Davis' account of expression.[1] According to Davis, a person expresses a mental state M (in the speaker sense of expression) by performing some observable act as an indication of M without thereby covertly simulating an unintentional indication of M. But, Green objects, an agent might satisfy these conditions and yet fail to express her M. Consider, for instance, a solo painter who tries to express her desolation but, perhaps due to the poor lighting in her basement, ends up painting a pastel-ridden picture of serenity. She paints the picture as an indication of her desolation, and so she satisfies Davis' conditions for expression; but, Green insists, she does not express her desolation.

What about my account? Can I make sense of her failure to express her desolation? There are actually a couple of things that can be said here. First of all, this seems to be precisely

---

[1] Green 2007b: 107.

the sort of case in which it becomes very easy to confuse one sense of expression with another. Wayne Davis, for instance, suggests that there is a sense of expression—which he calls the *evidential* sense of expression—in which a thing expresses a mental state by being evidence of that mental state.[2] So perhaps, because of its pastel color scheme, the painting fails to express desolation, in the sense that it fails to be evidence of the painter's desolation. But this is not yet to say that the painter does not a-express her desolation. And in fact, it's not really clear to me what it would even mean to say such a thing. Is her failure to express her desolation simply a matter of her, perhaps, not *feeling as though* she's expressed her desolation? I mean, suppose she actually considers the pastel picture to be a kind of ironic expression of her desolation. Would it still be the case that she fails to express her desolation? To say that it would, I think, would only be to assume that her ability to express her desolation depends in part upon her (or the painting's) ability to communicate that desolation to others. But as I argued above, it would be a mistake to make this assumption, since expression is something different than communication. On the other hand, maybe her failure to express really *is* just a matter of her seeing the painting in a better light and feeling as though she's failed to express her desolation. But why think a thing like that? Green certainly doesn't think that a speaker's ability to express a mental state depends in any way upon the speaker's feeling as though she's expressed that state. And rightly so, I think. While my account does make it, in some sense, up to the speaker whether or not she is expressing M by φing, it is not simply a matter of her feeling as though she hasn't expressed M that makes it the case that she hasn't, in fact, done so.

Having said all of that, though, I think my analysis of a-expression does leave enough room to account for the painter's failure to express her desolation. There are, as far as I can tell,

---

[2] Davis 2003: 43. Davis distinguishes *evidential* expression from two other senses of expression, which seem to me to be the same as Bar-On and Chrisman's *action* and *semantic* senses of expression.

at least a few ways of reading the condition that S φies intentionally, and the ambiguity here has to do with the various levels of generality at which we can describe actions. At a relatively general level, we might describe the painter's action as "painting a picture," and at that level, it would make sense to say that she φies intentionally. At a slightly less general level, though, we can describe her action as "painting a pastel picture." And at an even less general level, we can describe her action as "painting a pastel-ridden picture of serenity." But, you might think, since she didn't mean to paint a pastel-ridden picture of serenity, or even just a pastel picture, these are levels of description at which it would *not* be right to say that she φies intentionally.[3] So, at these levels of description, the painter fails to meet one of my conditions for expression, and so she fails to express her desolation. How, then, do we determine which level of description is the one at which we ought to be determining whether or not a speaker expresses her M by intentionally φing? This is actually a really complicated issue, one that I shouldn't (and don't) expect to settle here. But for now, I think it is at least plausible that, since it is the agent herself who is supposed to be expressing her own mental state by φing intentionally, the generality or specificity with which we describe her action ought to pay some respect to the generality or specificity of the content of her own intentions. Presumably, for instance, the content of the desolate painter's intention was something more specific than just to paint any ol' picture. And once we describe her action in a correspondingly specific sort of way, it will turn out, I'm betting, that she does not φ intentionally. And so, she fails to a-express her desolation because she fails to satisfy one of my conditions for expression.

---

[3] Here I assume the so-called "simple view" of intentional action according to which S does not φ intention-ally unless S had an intention to φ. But this is by no means uncontroversial. For more on the issue, see Adams 1986, Bratman 1987, and Knobe 2006.

Another way in which an expressivist might respond to my argument is to take issue with my understanding of belief. Here is Blackburn raising the sort of concern that I have in mind:

> It is not that it will be wrong, at the end of the day, to say that persons so expressing themselves are voicing the *belief* that something is good or right. They will be doing that. But since our explanatory task embraced the moral proposition, adopting it as a primitive is not an option. It is no good saying that people so saying express the belief that something is good or right, without more to tell us what that belief is. We need a theory of content.[4]

You will recall that Blackburn is what I am calling a deflationary expressivist. By deflating the truth predicate, he takes himself to have earned the expressivist the right to talk of the truth or falsity of moral claims. But there is more to Blackburn's deflationism than just the introduction of a kind of cognitivism to the expressivist's repertoire. In his book *Ruling Passions*, Blackburn actually urges a kind of deflationary *package*—there is more to deflate than just the truth predicate. And sure enough, part of the deflationary package is a theory of the content of moral beliefs.

> So what is it to believe that something is good, wonder whether it is good, to deny that it is good, … ? In basic or typical cases:
>
> > believing that X is good or right is roughly having an appropriately favourable valuation of X;
> >
> > wondering whether X is good or right is wondering what to do/what to admire or value;
> >
> > denying that X is good or right is rejecting a favourable attitude to X; …
>
> Here the practical states on the right-hand side are voiced and discussed in terms of attitudes to the saying or thought on the left.[5]

So Blackburn might respond to my argument by saying something like, "Okay, sure. When we use moral language, our reasons for doing so might very well be our beliefs about the goodness

---

[4] Blackburn unpublished: 3, italics mine.

[5] Blackburn 1998: 69-70.

or badness of things. And so, at least sometimes, we express beliefs when we use moral language. But these so-called 'beliefs' don't amount to anything above and beyond our conative attitudes toward things. So where's the problem for the expressivist?"

I have two things to say in response to this. First, it seems strange to have a theory of belief-content that applies only to some beliefs and not others. Why, for instance, should anyone but an expressivist agree that moral beliefs actually have an attitudinal content, while beliefs about, say, the shapes of things have exactly the sort of cognitive, propositional content that we ordinarily take such beliefs to have (which, I take it, Blackburn will still want to say)? In other words, you might think that beliefs ought to be *of a kind*, in the sense that they all have the same kind of content. And in fact, you might think that the expressivist could still be entitled to a legitimate form of cognitivism *without* having to adopt this heterogenous account of belief-contents. Some, for instance, distinguish between *thin* and *thick* senses of truth-aptness. A claim is thinly truth-apt if its syntactic form is such as to admit of ascriptions of truth or falsity, i.e., if the result of adding "It is true that …" or "… is true" is still grammatical. And a claim is thickly truth-apt if its content corresponds to some sort of appropriately mind-independent state of affairs. So "The mass breeding of dogs is morally wrong" is at least thinly truth-apt; and we can say this without having to say anything about the sort of mental state a speaker expresses when he makes such a claim, and especially, what kind of content that mental state has. In other words, a deflationary expressivist can have his cognitivism without making any controversial claims to the effect that some beliefs have attitudinal contents and others don't, i.e., he need not accept all of Blackburn's deflationary package, and he may in fact be better off that way.

A second problem, related to the first, is that Blackburn's theory of the content of moral beliefs seems to straightforwardly conflict with the way in which people ordinarily understand the nature of their moral beliefs, or at least, many of their moral beliefs. These sorts of issues

have only recently begun to spark the interests of psychologists and experimental philosophers, but one of the few studies to have been done so far suggests that ordinary people understand the contents of their moral beliefs to enjoy the same sort of objectivity as other sorts of non-moral beliefs, such as their beliefs about geographical and scientific matters.[6]

> The second major finding was that ethical beliefs were treated almost as objectively as scientific or factual beliefs, and decidedly more objectively than social conventions or tastes. … Arguably, many of our participants viewed their ethical beliefs as true in a mind-independent way. Such a view … implies that there is an objectively true fact of the matter concerning whether ethical beliefs are true.[7]

In other words, even if people do not ordinarily reflect upon the kind of content that their moral beliefs possess, they do ordinarily *treat* their moral beliefs in such a way as to suggest that they take those beliefs to have non-attitudinal contents. So by urging such a theory of the content of moral beliefs, Blackburn risks committing us all to a systematic confusion about the contents of our own mental states. And again, if one of the motivations for expressivism in the first place was to avoid an error theory, then committing competent users of moral language to an error with respect to the nature of their own beliefs is probably not the best route for the expressivist to take.

Finally, the expressivist may insist that I haven't taken his internalism seriously enough. My tripartite analysis of expression and subsequent argument against expressivism both assume that a speaker's beliefs can be among her reasons for action, but, the expressivist might say, this is just wrong. A mental state can only be a reason for S to φ if it is the sort of mental state that is capable of motivating S to φ. Or so the internalist about practical reasons insists. And since beliefs are not motivational states of mind, they cannot ever be among S's reasons for φing.

---

[6] Goodwin and Darley 2008.

[7] Goodwin and Darley 2008: 1359.

While I certainly do recognize this as a concern that my account of expression will have to address, the last thing I want to do at this point is get myself too entangled in the debate between internalists and externalists about reasons. So I'll only make a few brief comments in response. First of all, it is still very far from being a settled matter whether or not a consideration really does have to be capable of motivating S to φ in order for it to count as a reason for S to φ. Internalism about reasons may be a comfortable fit for the expressivist, but it is no less controversial a view for this. There is plenty to be said in favor of externalism, and indeed, it could turn out that the expressivist would be better off without being an internalist (doubtful, but possible). Second, the internalist might actually be talking past me here. There seems to me to be a clear difference between a thing's being able to *justify* S's φing and a thing's being able to *motivate* S to φ. Consider again what Blackburn himself has to say about some sets of attitudes being better than others. If there is an inconsistency in my own set of attitudes—e.g., suppose I've heretofore had a positive attitude toward physical health but a negative attitude toward physical exercise—then, presumably, I am *ipso facto* justified in revising my attitude set even if I am not thereby motivated to do so. When I say that M must be among S's reasons for φing, I mean only that M is able to justify (or play some part in justifying) S's φing. It is a different matter altogether whether or not M is able to motivate. Now, of course, the internalist may want to resist separating justification and motivation like this, but then how, for instance, would Blackburn ground his suggestion that I *ought* to revise my inconsistent set of attitudes regardless of whether I am, or could be, motivated to do so? And finally, we shouldn't ignore the difference between (a) *x*'s *being* among S's reasons for φing and (b) S *taking x to be* among her reasons to φ. Even if the internalist is right and S's beliefs really cannot ever be among her reasons to φ, there is still a problem for expressivists if S takes her beliefs to be among her

reasons for φing. For then they'd still have to say of S that she is confused in a way that we might find rather troubling.

7. CONCLUSION

In this paper, I've attempted to cast some light upon the concept of expression, and in particular, upon the sense of expression that seems to be lying at the heart of the EXPRESSIVIST THESIS, i.e., a-expression. If I am right, the relevant sense of expression is one according to which a speaker a-expresses a mental state by performing some speech act if, and only if, (a) the speaker has the mental state to express, (b) the speaker owns her speech act, and (c) the mental state is among the speaker's reasons for performing the speech act. The problem, though, is that, on any occasion in which a speaker recognizes some belief of hers as being among her reasons for uttering a moral sentence, then one of the following has to be the case: either (a) she does *not* express that belief, in which case the expressivist seems committed to the unwelcome conclusion that we are all often in error with respect to the nature of our own mental states and reasons for action; or (b) she *does* express that belief, in which case the EXPRESSIVIST THESIS is false.[1] If this dilemma is on point, I take it to provide us with a good reason for avoiding this particular branch of the flowchart I described in my introduction, i.e., to either accept something like Mackie's error theory, or to be a moral realist of some sort.

---

[1] Or she is misusing the language, but, you'll recall, we set aside the possibility of linguistic incompetence at the end of section 2.

REFERENCES

ADAMS, FREDERICK (1986). "Intention and Intentional Action: The Simple View," *Mind & Language* 1: 281-301.

AYER, A. J. (1952). *Language, Truth, and Logic*. New York: Dover.

BAR-ON, DORIT (1995). "'Meaning' Reconstructed: Grice and the Naturalization of Semantics," *Pacific Philosophical Quarterly* 76: 83-116.

———— (2004). *Speaking My Mind: Expression and Self-Knowledge*. Oxford: Clarendon Press.

BAR-ON, DORIT AND MATTHEW CHRISMAN (forthcoming). "Ethical Neo-Expressivism," forthcoming in *Oxford Studies in Metaethics, Vol. 4*. Available online at http://www.unc.edu/ ~dbar/papers/Bar-On&Chrisman_EthicalNeoExpressivism.pdf (last accessed March 21, 2009).

BLACKBURN, SIMON (1984). *Spreading the Word*. Oxford: Clarendon Press.

———— (1993). *Essays in Quasi-Realism*. Oxford: Oxford University Press.

———— (1998). *Ruling Passions*. Oxford: Clarendon Press.

———— (2007) "Pragmatism: All or Some?" Presented at the "Expressivism, Pragmatism, and Rep-resentationalism" conference, August 29-31, 2007, at the University of Sydney. Available online at http://www.phil.cam.ac.uk/~swb24/PAPERS/Representation.htm (last accessed March 21, 2009).

———— (unpublished). "Social and Individual Expression," available online at http://www.phil.cam. ac.uk/~swb24/PAPERS/Expression.pdf (last accessed March 21, 2009).

BOISVERT, DANIEL (2008). "Expressive-Assertivism," *Pacific Philosophical Quarterly* 89(2): 169-203.

BRATMAN, MICHAEL (1987). *Intention, Plans, and Practical Reason*. Cambridge: Harvard University Press.

COPP, DAVID (2001). "Realist-Expressivism: A Neglected Option for Moral Realism," *Social Philosophy and Policy* 18(2): 1-43.

DAVIS, WAYNE (2003). *Meaning, Expression, and Thought*. Cambridge: Cambridge University Press.

DIVERS, JOHN AND ALEXANDER MILLER (1994). "Why Expressivists About Value Should Not Love Minimalism About Truth," *Analysis* 54: 12-19.

DORR, CIAN (2002). "Non-Cognitivism and Wishful Thinking," *Noûs* 36(1): 97-103.

DRETSKE, FRED (1995). *Naturalizing the Mind*. Cambridge: MIT Press.

EILAN, NAOMI AND JOHANNES ROESSLER (2003). "Agency and Self-Awareness: Mechanisms and Epistemology," in Eilan and Roessler (eds.) *Agency and Self-Awareness: Issues in Philosophy and Psychology* (Oxford: Clarendon Press, 2003): 1-47.

GEACH, PETER (1965). "Assertion," *Philosophical Review* 74: 449-465.

GIBBARD, ALLAN (1990). *Wise Choices, Apt Feelings*. Cambridge: Harvard University Press.

GOODWIN, GEOFFREY P. AND JOHN M. DARLEY (2008). "The Psychology of Meta-Ethics: Exploring Objectivism," *Cognition* 106: 1339-1366.

GREEN, MITCHELL (2007a). "How Do Speech Acts Express Psychological States?" in Savas L. Tsohatzidis (ed.) *John Searle's Philosophy of Language: Force, Meaning and Mind* (Cambridge: Cambridge University Press, 2007): 267-84.

——— (2007b). *Self-Expression*. Oxford: Clarendon Press.

GRICE, H. P. (1957). "Meaning," *The Philosophical Review* 66(3): 377-388.

JACKSON, FRANK AND PHILIP PETTIT (1998). "A Problem for Expressivism," *Analysis* 58(4): 239-51.

KNOBE, JOSHUA (2006). "The Concept of Intentional Action: A Case Study in the Uses of Folk Psychology," *Philosophical Studies* 130: 203-231.

MARINO, PATRICIA (2005). "Expressivism, Deflationism, and Correspondence," *The Journal of Moral Philosophy* 2(2): 171-191.

O'BRIEN, LUCY (2003). "On Knowing One's Own Action," in Eilan and Roessler (eds.) *Agency and Self-Awareness: Issues in Philosophy and Psychology* (Oxford: Clarendon Press, 2003): 358-382.

ROESSLER, JOHANNES (2003). "Intentional Action and Self-Awareness," in Eilan and Roessler (eds.) *Agency and Self-Awareness: Issues in Philosophy and Psychology* (Oxford: Clarendon Press, 2003): 383-406.

SCHROEDER, MARK (2008a). *Being For*. Oxford: Clarendon Press.

——— (2008b). "Expression for Expressivists," *Philosophy and Phenomenological Research* 76(1): 86-116.

——— (2008c). "What is the Frege-Geach Problem?" *Philosophy Compass* 3(4): 703-720.

——— (forthcoming). "Hybrid Expressivism: Virtues and Vices," forthcoming in *Ethics*. Available online at http://www-rcf.usc.edu/~maschroe/research/Schroeder%20Hybrid.pdf.

SEARLE, JOHN (1969). *Speech Acts: An Essay in the Philosophy of Language*. Cambridge: Cambridge University Press.

SELLARS, WILFRID (1969). "Language as Thought and as Communication," *Philosophy and Phenomenological Research* 29: 506-527.

SMITH, MICHAEL (1994). "Why Expressivists About Value Should Love Minimalism About Truth," *Analysis* 54: 1-12.

STURGEON, NICHOLAS (1992). "Nonmoral Explanations," *Philosophical Perspectives* 6: 97-109.